

ABSTRACT OF THE DISCLOSURE

The present invention provides a system and method for comparing data sets, to ensure that they are accurate reflections of each other, without the need for performing $O(N^2)$ operations, in which N is the size of each data set. A hash table is generated for 5 the first data set. For each of the second data set entries, should the entry not exist in the hash table, the entry is second data set unique. Otherwise, the entry is removed from the hash table. At the end of the pass through the second data set entries, only those entries that exist in the hash table are first data set unique. Alternately, two processes operate in parallel so that each selects entries from one of the data sets and determines if the entry 10 exists in the hash table. If the entry does exist, it is removed. Otherwise, the entry is added to the hash table.